This protein has a molecular weight of 78 to 81 kDa, as measured by SDS-PAGE. For purposes of the present application, the term "thermostable" refers to a DNA ligase which is resistant to inactivation by heat.

The thermostable ligase of the present invention has a 100 fold higher fidelity than T4 ligase and 6 fold higher fidelity than wild-type *Thermus thermophilus* ligase, when sealing a ligation junction between a pair of oligonucleotide probes hybridized to a target sequence where there is a mismatch with the oligonucleotide probe having its 3' end abutting the ligation junction at the base immediately adjacent the ligation junction. This ligase also has a 50 fold higher fidelity than T4 ligase and 5 fold higher fidelity than wild-type *Thermus thermophilus* ligase, when sealing a ligation junction between a pair of oligonucleotide probes hybridized to a target sequence where there is a mismatch with the oligonucleotide probe having its 3' end abutting the ligation junction at the base penultimate to the ligation junction. Finally, the thermostable ligase of the present invention, in the presence of a Mn^{2+} cofactor, has a 12 fold higher fidelity than wild-type *Thermus thermophilus* ligase, when sealing a ligation junction between a pair of oligonucleotide probes hybridized to a target sequence where there is a mismatch with the oligonucleotide probe having its 3' end abutting the ligation junction at the base immediately adjacent to the ligation junction. For purposes of the present invention, "fidelity" is defined to mean the ratio of the initial rate of ligating two adjacent probes hybridized to a complementary template with a C-G match at the base of the probe with its 3' end at the ligation junction to the initial rate of ligating two adjacent probes hybridized to a complementary template with a G-T mismatch at the base of the probe with its 3' end at the ligation junction.

The thermostable ligase of the present invention is also characterized by having an arginine adjacent to the active site lysine (i.e. K) in the KXDG motif (where X is any amino acid).

This protein is encoded by a DNA molecule having a nucleotide sequence of SEQ. ID. No. 2 as follows:

ATGACCCTAGAGGAGGCCCGCAGGCGCGTCAACGAACTCAGGGACCTGATCCGTTAC
CACAACTACCTCTATTACGTCTTGGACGCCCCCGAGATCTCCGACGCCGAGTACGAC
CGGCTCCTTAGGGAGCTTAAGGAGCTGGAGGAGCGCTTTCCCGAGCTCAAAAGCCCC

GACTCCCCCACGGAACAGGTGGGGGCGAGGCCTCTGGAGGCCACCTTCCGCCCGGTG
CGCCACCCCACCCGCATGTACTCCCTGGACAACGCCTTTTCCTTGGACGAGGTGAGG
GCCTTTGAGGAGCGCATAGAGCGGGCCCTGGGGCGGAAGGGGCCCTTCCTCTACACC
GTGGAGCGCAAGGTGGACGGTCTTTCCGTGAACCTCTACTACGAGGAGGGCATCCTC
GTCTTTGGGGCCACCCGGGGCGACGGGGAGACCGGGGAGGAGGTGACCCAGAACCTC
CTCACCATCCCCACCATTCCCCGCCGCCTCACGGGCGTTCCGGACCGCCTCGAGGTC
CGGGGCGAGGTCTACATGCCCATAGAGGCCTTCCTCAGGCTCAACCAGGAGCTGGAG
GAGGCGGGGGAGCGCATCTTCAAAAACCCCAGGAACGCCGCCGCCGGGTCCTTGCGG
CAGAAAGACCCCAGGGTCACGGCCAGGCGGGGCCTGAGGGCCACCTTTTACGCCCTG
GGGCTGGGCCTGGAGGAAACCGGGTTAAAAAGCCAGCACGACCTTCTCCTATGGCTA
AGAGAGCGGGGCTTTCCCGTGGAGCACGGCTTTACCCGGGCCCTGGGGGCGGAGGGG
GTGGAGGAGGTCTACCAGGCCTGGCTCAAGGAGAGGCGGAAGCTTCCCTTTGAGGCC
GACGGGGTGGTGGTCAAGCTGGACGACCTCGCCCTCTGGCGGGAGCTGGGGTACACC
GCCCGCACCCCCCGCTTCGCCCTCGCCTACAAGTTCCCGGCCGAGGAGAAGGAGACC
CGCCTCCTCTCCGTGGCCTTCCAGGTGGGGCGGACCGGGCGCATCACCCCCGTGGGC
GTTCTGGAGCCCGTCTTCATAGAGGGCAGCGAGGTGAGCCGGGTCACCCTCCACAAC
GAGAGCTTCATTGAGGAGCTGGACGTGCGCATCGGCGACTGGGTGCTGGTCCACAAG
GCGGGCGGGGTGATTCCCGAGGTGCTGAGGGTCCTGAAAGAGCGCCGCACCGGGGAG
GAGAAGCCCATCATCTGGCCCGAGAACTGCCCCGAGTGCGGCCACGCCCTCATCAAG
GAGGGGAAGGTCCACCGCTGCCCCAACCCCTTGTGCCCCGCCAAGCGCTTTGAGGCC
ATCCGCCACTACGCCTCCCGCAAGGCCATGGACATCCAGGGCCTGGGGGAGAAGCTC
ATAGAAAAGCTTCTGGAAAAGGGCCTGGTCCGGGACGTGGCCGACCTCTACCGCCTG
AAGAAGGAGGACCTGGTGAACCTGGAGCGCATGGGGGAGAAGAGCGCAGAGAACCTC
CTCCGCCAGATAGAGGAGAGCAAGGGCCGCGGCCTGGAGCGCCTCCTTTACGCCCTG
GGCCTTCCCGGGGTGGGGGAGGTGCTGGCCCGGAACCTGGCCCTCCGCTTCGGCCAC
ATGGACCGCCTTCTGGAGGCGGGCCTCGAGGACCTCCTGGAGGTGGAGGGGGTGGGC
GAGCTCACCGCCCGGGCCATCCTGAATACCCTAAAGGACCCGGAGTTCCGGGACCTG
GTGCGCCGCCTGAAGGAGGCCGGGGTGGAGATGGAGGCCAAAGAGCGGGAGGGCGAG
GCCTTGAAGGGGCTCACCTTCGTCATCACCGGGGAGCTTTCCCGGCCCCGGGAGGAG
GTGAAGGCCCTCCTTAGGCGGCTTGGGGCCAAGGTGACGGACTCGGTGAGCCGCAAG
ACGAGCTTCCTGGTGGTGGGGGAGAACCCGGGGAGCAAGCTGGAAAAGGCCCGCGCC
TTGGGGGTCCCCACCCTGAGCGAGGAGGAGCTCTACCGCCTCATTGAGGAGAGGACG
GGCAAGGACCCAAGGGCCCTCACGGCCTAG

Fragments of the above polypeptide or protein are also encompassed by the present invention.

Suitable fragments can be produced by several means. In the first, subclones of the gene encoding the protein of the present invention are produced by conventional molecular genetic manipulation by subcloning gene fragments. The subclones then are expressed *in vitro* or *in vivo* in bacterial cells to yield a smaller protein or peptide that can be tested for ligase activity according to the procedure described below.

As an alternative, fragments of the ligase of the present invention can be produced by digestion of the full-length ligase with proteolytic enzymes like

chymotrypsin or *Staphylococcus* proteinase A, or trypsin. Different proteolytic enzymes are likely to cleave ligase proteins at different sites based on the amino acid sequence of the ligase. Some of the fragments that result from proteolysis may be active ligases.

In another approach, based on knowledge of the primary structure of the protein, fragments of the ligase encoding gene may be synthesized by using the PCR technique together with specific sets of primers chosen to represent particular portions of the protein. These then would be cloned into an appropriate vector for increased expression of a truncated peptide or protein.

Chemical synthesis can also be used to make suitable fragments. Such a synthesis is carried out using known amino acid sequences for the ligase being produced. Alternatively, subjecting the full length ligase to high temperatures and pressures will produce fragments. These fragments can then be separated by conventional procedures (e.g., chromatography, SDS-PAGE).

Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the properties, secondary structure and hydropathic nature of the polypeptide. For example, a polypeptide may be conjugated to a signal (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification, or identification of the polypeptide.

Suitable DNA molecules are those that hybridize to a DNA molecule comprising a nucleotide sequence of 50 continuous bases of SEQ. ID. No. 2 under stringent conditions characterized by a hybridization buffer comprising 0.9M sodium citrate ("SSC") buffer at a temperature of 37°C and remaining bound when subject to washing with the SSC buffer at 37°C; and preferably in a hybridization buffer comprising 20% formamide in 0.9M saline/0.09M SSC buffer at a temperature of 42°C and remaining bound when subject to washing at 42°C with 0.2x SSC buffer at 42°C.

The protein or polypeptide of the present invention is preferably produced in purified form (preferably at least about 80%, more preferably 90%, pure) by conventional techniques. Typically, the protein or polypeptide of the present